Static Program Analysis

Lecture 6: Dataflow Analysis V (MOP vs. Fixpoint Solution)

Thomas Noll

Lehrstuhl für Informatik 2
(Software Modeling and Verification)

RWTH AACHEN UNIVERSITY

noll@cs.rwth-aachen.de

http://moves.rwth-aachen.de/teaching/ws-1415/spa/

Winter Semester 2014/15

Outline

The MOP Solution I

- Other solution method for dataflow systems
- MOP = Meet Over all Paths
- Analysis information for block B^l
  = least upper bound over all paths leading to l
  = most precise information for l ("reference solution")

Definition (Paths)

Let $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a dataflow system. For every $l \in Lab$, the set of paths up to l is given by

$$Path(l) := \{[l_1, \ldots, l_{k-1}] \mid k \geq 1, l_1 \in E, \\ (l_i, l_{i+1}) \in F \text{ for every } 1 \leq i < k, l_k = l\}.$$

For a path $\pi = [l_1, \ldots, l_{k-1}] \in Path(l)$, we define the transfer function $\varphi_\pi : D \to D$ by

$$\varphi_\pi := \varphi_{l_{k-1}} \circ \ldots \circ \varphi_{l_1} \circ \mathrm{id}_D$$

(so that $\varphi_{[]} = \mathrm{id}_D$).

The MOP Solution II

Definition (MOP solution)

Let $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a dataflow system where $Lab = \{l_1, \ldots, l_n\}$. The MOP solution for S is determined by

$$\mathrm{mop}(S) := (\mathrm{mop}(l_1), \ldots, \mathrm{mop}(l_n)) \in D^n$$

where, for every $l \in Lab$,

$$\mathrm{mop}(l) := \bigsqcup\{\varphi_\pi(\iota) \mid \pi \in Path(l)\}.$$

Remark:

- $Path(l)$ is generally infinite

$\Longrightarrow$ not clear how to compute $\mathrm{mop}(l)$

- In fact: MOP solution generally undecidable (later)

Outline

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of Constant Propagation Analysis is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of Constant Propagation Analysis is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for Constant Folding: replace reference to variable by constant value and evaluate constant expressions

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of Constant Propagation Analysis is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for Constant Folding: replace reference to variable by constant value and evaluate constant expressions

Example (Constant Propagation Analysis)

```
[x := 1]^1;
[y := 1]^2;
[z := 1]^3;
while [z > 0]^4 do
  [w := x+y]^5;
  if [w = 2]^6 then
    [x := y+2]^7
```

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of Constant Propagation Analysis is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for Constant Folding: replace reference to variable by constant value and evaluate constant expressions

Example (Constant Propagation Analysis)

```
[x := 1]^1;
[y := 1]^2;
[z := 1]^3;
while [z > 0]^4 do
  [w := x+y]^5;
  if [w = 2]^6 then
    [x := y+2]^7
```

- $y = z = 1$ at labels 4–7

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of Constant Propagation Analysis is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for Constant Folding: replace reference to variable by constant value and evaluate constant expressions

Example (Constant Propagation Analysis)

```
[x := 1]^1;
[y := 1]^2;
[z := 1]^3;
while [z > 0]^4 do
  [w := x+y]^5;
  if [w = 2]^6 then
    [x := y+2]^7
```

- y = z = 1 at labels 4–7
- w, x not constant at labels 4–7

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of Constant Propagation Analysis is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for Constant Folding: replace reference to variable by constant value and evaluate constant expressions

Example (Constant Propagation Analysis)

```
[x := 1]^1;
[y := 1]^2;
[z := 1]^3;
while [z > 0]^4 do
  [w := x+y]^5;
  if [w = 2]^6 then
    [x := y+2]^7
```

- y = z = 1 at labels 4–7
- w, x not constant at labels 4–7
- possible optimisations:
  `[true]^4 [w := x+1]^5 [x := 3]^7`

Formalising Constant Propagation Analysis I

The dataflow system $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ is given by

- set of labels $Lab := Lab_c$,
- extremal labels $E := \{\text{init}(c)\}$ (forward problem),
- flow relation $F := \text{flow}(c)$ (forward problem),
- complete lattice $(D, \sqsubseteq)$ where
  - $D := \{\delta \mid \delta : Var_c \to \mathbb{Z} \cup \{\bot, \top\}\}$
    - $\delta(x) = z \in \mathbb{Z}$: x has constant value z
    - $\delta(x) = \bot$: x undefined
    - $\delta(x) = \top$: x overdefined (i.e., several possible values)
  - $\sqsubseteq \subseteq D \times D$ defined by pointwise extension of $\bot \sqsubseteq z \sqsubseteq \top$ (for every $z \in \mathbb{Z}$)

Example

$Var_c = \{\mathtt{w}, \mathtt{x}, \mathtt{y}, \mathtt{z}\}$,

$\delta_1 = (\underbrace{\bot}_{\mathtt{w}}, \underbrace{1}_{\mathtt{x}}, \underbrace{2}_{\mathtt{y}}, \underbrace{\top}_{\mathtt{z}}),\ \delta_2 = (\underbrace{3}_{\mathtt{w}}, \underbrace{1}_{\mathtt{x}}, \underbrace{4}_{\mathtt{y}}, \underbrace{\top}_{\mathtt{z}})$

$\Longrightarrow\ \delta_1 \sqcup \delta_2 = (\underbrace{3}_{\mathtt{w}}, \underbrace{1}_{\mathtt{x}}, \underbrace{\top}_{\mathtt{y}}, \underbrace{\top}_{\mathtt{z}})$

Formalising Constant Propagation Analysis II

Dataflow system $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ (continued):

- extremal value $\iota := \delta_\top \in D$ where $\delta_\top(x) := \top$ for every $x \in Var_c$ (i.e., every x has (unknown) default value)
- transfer functions $\{\varphi_l \mid l \in Lab\}$ defined by

$$\varphi_l(\delta) := \begin{cases} \delta & \text{if } B^l = \texttt{skip} \text{ or } B^l \in BExp \\ \delta[x \mapsto val_\delta(a)] & \text{if } B^l = (x \texttt{ := } a) \end{cases}$$

where

$$val_\delta(x) := \delta(x) \qquad val_\delta(a_1 \; op \; a_2) := \begin{cases} z_1 \; op \; z_2 & \text{if } z_1, z_2 \in \mathbb{Z} \\ \bot & \text{if } z_1 = \bot \text{ or } z_2 = \bot \\ \top & \text{otherwise} \end{cases}$$
$$val_\delta(z) := z$$

for $z_1 := val_\delta(a_1)$ and $z_2 := val_\delta(a_2)$

Formalising Constant Propagation Analysis III

Example

If $\delta = (\underbrace{\bot}_{\mathtt{w}}, \underbrace{1}_{\mathtt{x}}, \underbrace{2}_{\mathtt{y}}, \underbrace{\top}_{\mathtt{z}})$, then

$$\varphi_l(\delta) = \begin{cases} (\underbrace{0}_{\mathtt{w}}, \underbrace{1}_{\mathtt{x}}, \underbrace{2}_{\mathtt{y}}, \underbrace{\top}_{\mathtt{z}}) & \text{if } B^l = (\mathtt{w\ :=\ 0}) \\ (\underbrace{3}_{\mathtt{w}}, \underbrace{1}_{\mathtt{x}}, \underbrace{2}_{\mathtt{y}}, \underbrace{\top}_{\mathtt{z}}) & \text{if } B^l = (\mathtt{w\ :=\ y+1}) \\ (\underbrace{\bot}_{\mathtt{w}}, \underbrace{1}_{\mathtt{x}}, \underbrace{2}_{\mathtt{y}}, \underbrace{\top}_{\mathtt{z}}) & \text{if } B^l = (\mathtt{w\ :=\ w+x}) \\ (\underbrace{\top}_{\mathtt{w}}, \underbrace{1}_{\mathtt{x}}, \underbrace{2}_{\mathtt{y}}, \underbrace{\top}_{\mathtt{z}}) & \text{if } B^l = (\mathtt{w\ :=\ z+2}) \end{cases}$$

Outline

Example of Constant Propagation Analysis

Example 6.1

Constant Propagation Analysis for

$$
\begin{array}{ll}
c := [\texttt{x := 1}]^1; & \varphi_1(a,b,c,d) = (a,1,c,d) \\
\quad [\texttt{y := 1}]^2; & \varphi_2(a,b,c,d) = (a,b,1,d) \\
\quad [\texttt{z := 1}]^3; & \varphi_3(a,b,c,d) = (a,b,c,1) \\
\quad \texttt{while } [\texttt{z > 0}]^4 \texttt{ do} & \varphi_4(a,b,c,d) = (a,b,c,d) \\
\qquad [\texttt{w := x+y}]^5; & \varphi_5(a,b,c,d) = (b+c,b,c,d) \\
\qquad \texttt{if } [\texttt{w = 2}]^6 \texttt{ then} & \varphi_6(a,b,c,d) = (a,b,c,d) \\
\qquad\quad [\texttt{x := y+2}]^7 & \varphi_7(a,b,c,d) = (a,c+2,c,d)
\end{array}
$$

(for $\delta = (\delta(\mathtt{w}), \delta(\mathtt{x}), \delta(\mathtt{y}), \delta(\mathtt{z})) \in D$)

1. Fixpoint solution (on the board)
2. MOP solution (on the board)

Outline

MOP vs. Fixpoint Solution I

Example 6.2 (Constant Propagation)

$c :=$
```
if [z > 0]^1 then
    [x := 2;]^2
    [y := 3;]^3
else
    [x := 3;]^4
    [y := 2;]^5
[z := x+y;]^6
[...]^7
```

MOP vs. Fixpoint Solution I

Example 6.2 (Constant Propagation)

$c :=$ `if [z > 0]`^{1} `then`
`[x := 2;]`^{2}
`[y := 3;]`^{3}
`else`
`[x := 3;]`^{4}
`[y := 2;]`^{5}
`[z := x+y;]`^{6}
$[\ldots]^7$

Transfer functions
(for $\delta = (\delta(\mathtt{x}), \delta(\mathtt{y}), \delta(\mathtt{z})) \in D$):

$$\varphi_1(a, b, c) = (a, b, c)$$
$$\varphi_2(a, b, c) = (2, b, c)$$
$$\varphi_3(a, b, c) = (a, 3, c)$$
$$\varphi_4(a, b, c) = (3, b, c)$$
$$\varphi_5(a, b, c) = (a, 2, c)$$
$$\varphi_6(a, b, c) = (a, b, a + b)$$

MOP vs. Fixpoint Solution I

Example 6.2 (Constant Propagation)

$c :=$ `if [z > 0]`^{1} `then`
    `[x := 2;]`^{2}
    `[y := 3;]`^{3}
  `else`
    `[x := 3;]`^{4}
    `[y := 2;]`^{5}
  `[z := x+y;]`^{6}
  $[\ldots]^7$

Transfer functions
(for $\delta = (\delta(\mathtt{x}), \delta(\mathtt{y}), \delta(\mathtt{z})) \in D$):

$$\begin{aligned}
\varphi_1(a,b,c) &= (a,b,c)\\
\varphi_2(a,b,c) &= (2,b,c)\\
\varphi_3(a,b,c) &= (a,3,c)\\
\varphi_4(a,b,c) &= (3,b,c)\\
\varphi_5(a,b,c) &= (a,2,c)\\
\varphi_6(a,b,c) &= (a,b,a+b)
\end{aligned}$$

1. Fixpoint solution:

$$\begin{aligned}
\mathrm{CP}_1 &= \iota &&= (\top,\top,\top)\\
\mathrm{CP}_2 &= \varphi_1(\mathrm{CP}_1) &&= (\top,\top,\top)\\
\mathrm{CP}_3 &= \varphi_2(\mathrm{CP}_2) &&= (2,\top,\top)\\
\mathrm{CP}_4 &= \varphi_1(\mathrm{CP}_1) &&= (\top,\top,\top)\\
\mathrm{CP}_5 &= \varphi_4(\mathrm{CP}_4) &&= (3,\top,\top)\\
\mathrm{CP}_6 &= \varphi_3(\mathrm{CP}_3) \sqcup \varphi_5(\mathrm{CP}_5) &&\\
&= (2,3,\top) \sqcup (3,2,\top) &&= (\top,\top,\top)\\
\mathrm{CP}_7 &= \varphi_6(\mathrm{CP}_6) &&= (\top,\top,\top)
\end{aligned}$$

MOP vs. Fixpoint Solution I

Example 6.2 (Constant Propagation)

$c :=$ `if [z > 0]`^{1} `then`

```
      [x := 2;]²
      [y := 3;]³
    else
      [x := 3;]⁴
      [y := 2;]⁵
    [z := x+y;]⁶
    [...]⁷
```

Transfer functions (for $\delta = (\delta(\mathtt{x}), \delta(\mathtt{y}), \delta(\mathtt{z})) \in D$):

$$\begin{aligned}
\varphi_1(a,b,c) &= (a,b,c)\\
\varphi_2(a,b,c) &= (2,b,c)\\
\varphi_3(a,b,c) &= (a,3,c)\\
\varphi_4(a,b,c) &= (3,b,c)\\
\varphi_5(a,b,c) &= (a,2,c)\\
\varphi_6(a,b,c) &= (a,b,a+b)
\end{aligned}$$

1. Fixpoint solution:

$$\begin{aligned}
\mathrm{CP}_1 &= \iota && = (\top,\top,\top)\\
\mathrm{CP}_2 &= \varphi_1(\mathrm{CP}_1) && = (\top,\top,\top)\\
\mathrm{CP}_3 &= \varphi_2(\mathrm{CP}_2) && = (2,\top,\top)\\
\mathrm{CP}_4 &= \varphi_1(\mathrm{CP}_1) && = (\top,\top,\top)\\
\mathrm{CP}_5 &= \varphi_4(\mathrm{CP}_4) && = (3,\top,\top)\\
\mathrm{CP}_6 &= \varphi_3(\mathrm{CP}_3) \sqcup \varphi_5(\mathrm{CP}_5) &&\\
&= (2,3,\top) \sqcup (3,2,\top) && = (\top,\top,\top)\\
\mathrm{CP}_7 &= \varphi_6(\mathrm{CP}_6) && = (\top,\top,\top)
\end{aligned}$$

2. MOP solution:

$$\begin{aligned}
\mathrm{mop}(7) &= \varphi_{[1,2,3,6]}(\top,\top,\top) \sqcup\\
&\quad\ \varphi_{[1,4,5,6]}(\top,\top,\top)\\
&= (2,3,5) \sqcup (3,2,5)\\
&= (\top,\top,5)
\end{aligned}$$

MOP vs. Fixpoint Solution II

Theorem 6.3 (MOP vs. Fixpoint Solution)

Let $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ *be a dataflow system. Then*

$$\mathrm{mop}(S) \sqsubseteq \mathrm{fix}(\Phi_S)$$

Reminder: by Definition 4.9,

$$\Phi_S : D^n \to D^n : (d_1, \ldots, d_n) \mapsto (d'_1, \ldots, d'_n)$$

where $Lab = \{1, \ldots, n\}$ and, for each $l \in Lab$,

$$d'_l := \begin{cases} \iota & \text{if } l \in E \\ \bigsqcup\{\varphi_{l'}(d_{l'}) \mid (l', l) \in F\} & \text{otherwise} \end{cases}$$

MOP vs. Fixpoint Solution II

Theorem 6.3 (MOP vs. Fixpoint Solution)

Let $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ *be a dataflow system. Then*

$$\mathsf{mop}(S) \sqsubseteq \mathsf{fix}(\Phi_S)$$

Reminder: by Definition 4.9,

$$\Phi_S : D^n \to D^n : (d_1, \ldots, d_n) \mapsto (d'_1, \ldots, d'_n)$$

where $Lab = \{1, \ldots, n\}$ and, for each $l \in Lab$,

$$d'_l := \begin{cases} \iota & \text{if } l \in E \\ \bigsqcup\{\varphi_{l'}(d_{l'}) \mid (l', l) \in F\} & \text{otherwise} \end{cases}$$

Proof.

on the board □

Remark: as Example 6.2 shows, $\mathsf{mop}(S) \neq \mathsf{fix}(\Phi_S)$ is possible